EMBEDDED QUADTREE WAVELETS IN IMAGE COMPRESSION

Field of the Invention

5   The present invention relates generally to image coding, and more particularly to compression and decompression of digital images.

Background of the Invention

The advent of multimedia computing has created an increased demand for high-
10   performance image compression systems. In the last few years, the wavelet transform has become a mainstream, base technology for image compression coding. Wavelet transforms, otherwise known as hierarchical subband decompositions, result in multi-resolution decomposition hierarchy (MDH) representations of the source image as illustrated in Fig. 1. Bit rates lower than 1 bit/pixel can be achieved through the
15   efficient coding of the wavelet transform coefficients generated in the production of the MDH data.

A most important and beneficial characteristic of the wavelet coefficients generated by the transform is that most of the coefficients will possess very small amplitudes that
20   will reduce to zeros after scalar quantization. For many image processing purposes, the importance or significance of a wavelet transform coefficient can be measured by its absolute value in relation to predetermined threshold values. A wavelet coefficient is said to be significant or insignificant, in relation to a particular threshold value, depending on whether or not its magnitude exceeds that threshold. The importance of
25   a set of wavelet coefficients can be collectively ascertained using a "significance map". A "significance map" is a bitmap recording the location of the significant coefficients. A large fraction of the bit budget may be spent on encoding the significance map. Therefore, the compression performance of an image coding system largely relies on its efficiency in coding the significance map.

30

In U.S. Patent 5412741 J. M. Shapiro disclosed an embedded zerotree wavelet algorithm called "EZW". A more efficient implementation of this invention, called set

partitioning in hierarchical trees or "SPIHT" was disclosed by Said et al. in "A New, Fast, and Efficient Image Codec Based-on-Set Partitioning in Hierarchical Trees", A. Said and W. Pearlman, IEEE Trans. On Circuits and Systems For Video Technology, Vol. 6, No. 3, June, 1996.

5

Because of its inherent simplicity, efficiency and competitiveness in performance to most other techniques, EZW-based coding has been considered one of the best in the image compression research community. Further, it has been chosen as a candidate technique for the new generation International Standard for image (JPEG 2000) and

10   video (MPEG 4) coding.

EZW-based coding techniques consist of three basic methodological elements. The first element is the partial ordering of the MDH data by amplitude. By duplicating the ordering information at the decoder, such that the MDH data with larger amplitude

15   will be transmitted first, it is assured that the transformation coefficients carrying a larger amount of information will more probably be available in reconstructing the image. Usually, the partial ordering is performed using a set of octave decreasing thresholds. The second element is the ordered bit plane transmission of refinement bits in order to achieve the embedded quantization. The third element is to make use of the

20   cross, sub-band correlation between the amplitudes of MDH data to code the significance map.

Although the zerotree structure has proven successful in coding MDH data, it is not the only logical exploitation of the data set's inherent regularities. EZW is not the most

25   efficient representation when considering the compactness of the resulting code nor does the completely closed structure of the zerotree method allow for independent or parallel processing. In the case of a zerotree-coded, multi-layer representation of a visual object like an MPEG-4 object, only the base layer can be independently decoded. The decoding of all enhancement layers must rely on the information of

30   previously decoded layers. In other words, the zerotree representation of objects inherently prevents independent decodability. This inseparability also introduces a higher susceptibility to bit errors. A single bit error could potentially, after

interpretation at each succeeding resolution level, lead to decoder derailment. Finally, the closed structure of zerotree representation makes it difficult to add in new coding methods or features.

5   **Summary of the Invention**

The present invention is a method of compressing grayscale and color image data with a high degree of compression performance. An objective of the present invention is to provide a compressibly efficient, fast, method and system to code the significance information of the wavelet transform coefficients. A further objective is to provide a

10   method and system of producing a compressed bit-stream that is scalable, region-based accessible, robust to errors, and independently decodable. The present invention provides a logically simple and fast method of coding that possess and a high degree of parallelism that lends itself to hardware implementation. The bit-stream produced by the present system is more robust to bit error than the prior art since all sub-band

15   blocks are encoded independently and errors at one scale will not lead to errors in other scales.

In accordance with an aspect of the instant invention there is provided a method for encoding and decoding digital still images to produce a scalable, content accessible

20   compressed bit stream comprising the steps of decomposing and ordering the raw image data into a hierarchy of multi-resolution sub-images; setting an initial threshold of significance and creating a significance index; determining an initial list of insignificant blocks; forming the list of significant coefficients by encoding a significant map using a quadtree representation; recursively reducing the threshold values and

25   repeating the encoding process for each threshold value; and then transmitting refinement bits of significant coefficients.

In accordance with another aspect of the instant invention there is provided an apparatus for encoding and decoding of digital still images that produces a scalable,

30   content accessible compressed bit stream comprising a means of decomposing and ordering the raw image data into a hierarchy of multi-resolution sub-images; means for setting an initial threshold of significance and creating a significance index; means for

determining an initial list of insignificant blocks; means of forming the list of significant coefficients by encoding a significant map using a quadtree representation; a means of recursively reducing the threshold values and repeating the encoding process; and a means by which refinement bits of significant coefficients are transmitted.

5

In accordance with yet another aspect of the instant invention there is provided method of decoding digital still images to produce a scalable, content accessible compressed bit stream comprising the steps of: decoding the bitstream header; determining the initial threshold values and the array of initial significant pixels,

10   insignificant bits and wavelet coefficients; decoding the significance maps; modifying the significance lists and decoding the refinement bits for each threshold level; reconstruct the wavelet coefficient array; perform the inverse wavelet transform; and reconstructing the image.

15   **Brief Description of the Drawings**

Figure 1 is a schematic illustration of a three-layer wavelet decomposition.

Figure 1a is a graphic illustration of a three-layer wavelet decomposition performed on

20   the test image "Lena".

Figure 2 illustrates the binary representation of a wavelet transform coefficient after it is converted into an integer form.

25   Figure 3 is a block diagram of the invented image encoder.

Figure 4 is the process of initializing the lists LSP and LIB.

Figure 5 illustrates the algorithm that determines the initial threshold.

30

Figure 6 is a flowchart of the quadtree coding of the significance map.

Figure 7 is a flowchart of the refinement process.

Figure 8 is a block diagram of the multiplexer.

5   Figure 9 illustrates the default order of data packing.

Figure 10 is a block diagram of the image decoder of the invention.

Figure 11 is a flowchart of the quadtree decoding of the significance map.

10

Detailed Description of Preferred Embodiments

When the wavelet transform of a preferred embodiment is applied to decompose an image it results in four frequency sub-band signals. These sub-bands are: high horizontal, high vertical or "HH", high horizontal low vertical "HL", low horizontal

15   high vertical "LH", and low horizontal low vertical "LL", frequency sub-bands. The LL sub-band is then further wavelet-transformed to produce a further set of HH, HL, HL, and LL sub-bands. This procedure is performed recursively to produce a multi-resolution decomposition hierarchy (MDH) of the original image. This is illustrated in Fig.1 where three levels of transformation have been applied. Of course, the skilled

20   reader will appreciate that an arbitrary number of sub-band decompositions may be applied.

In Fig. 1 the lowest frequency sub-band i.e the sub-band that provides the coarsest resolution scale, is that at the top, left-most block 101 represented by LL3. The highest

25   frequency sub-bands or those at the finest resolution scale are the blocks HL1 102, LH1 103, and HH1 104.

Figure 1a is a graphic illustration of the present invention's three layer wavelet decomposition of the test image Lena. The original image 1a01 can be seen to have 3

30   levels of resolution in the decomposed image 1a02. The high frequency data of HH1 104 can be seen to offer the most detail in the bottom, right-most block 1a03.

After a wavelet transform has occurred, each pixel is represented by a wavelet transform coefficient. In the preferred embodiment of the current invention, each of these coefficients is represented in a fixed-point, binary format, most typically with less than 16 bits, and treated as an integer. Fig. 2 illustrates the binary representation in the

5   general case of a wavelet transform coefficient. In this system, the first bit 201 is dedicated to represent its sign -- positive or negative. The first non-zero bit 202 following the sign bit is called the leading one bit or LOB. The position of the LOB is determined by the magnitude of the coefficient. That is to say that the larger the value of the coefficient, the more closely after the sign bit will it occur. All of the bits

10   following the LOB 202 are called refinement bits 203.

After the coefficients are generated in the wavelet transformation and are given their binary representation, three lists are initialized. The first of these is called the list of significant pixels or LSP. Each entry in LSP corresponds to an individual pixel on the

15   MDH plane and is identified by a pair of coordinates (i,j). The LSP is initialized as an empty list since the significance of individual pixels has yet to be determined. The second list is called the list of insignificant blocks or LIB. The entries in this list are composed of the coordinates of the left-top pixel of a block of coordinates (i1,j1) plus the width and height of the block (i2,j2) measured in pixels. An entry in the LIB

20   represents a block made up of an individual pixel when $i2 = j2 = 1$. When first initialized the TLIB is empty. After the lists are initialized, each sub-band block becomes an entry in LIB. The order of the entries in the initial LIB can be arranged arbitrarily but the default order of sub-band entry is LL3, LH3, HL3, HH3, LH2, HL2, HH2, LH1, HL1, HH1. Figure 4 represents the decision tree for the creation of LSP

25   and the default entry into the LIB.

The next step in the formulation of the lists is the calculation of threshold values to determine the significance of the coefficients. After the wavelet transform, the maximum magnitude "M" of all transform coefficients must be determined. One skilled

30   in the art is familiar with the fact that the vast majority of coefficients from an efficiently implemented MDH will have relatively low values. Once M has been determined, a value N is found which satisfies the condition: $2^N \leq M < 2^{N+1}$. The initial

6

threshold is set at 2^N, and the set of various N values is called the threshold index. The threshold values then decrease by powers-of-2-for-ease of bit-wise computation. At each threshold value a significance map is produced by comparing the coefficients with the threshold value. Those coefficients that exceed the threshold are given a value of 1

5   and thus join the map of significant coefficients. Coefficients less than the threshold value are given a value of zero in that significant map. A significance map for each threshold value, in the form of a binary image, is thus produced.

Recalling that the LIB is first composed of the sub-band blocks of the MDH, the

10   preferred embodiment of the present invention, begins the quadtree encoding of the significance data. For the given block, we count the number of significant coefficients in this block. If the number is zero the identifying coordinates of this square are added to TLIB. If there is at least one significant coefficient in this block, "the parent block", it is divided into four equal-sized sub-blocks called "child blocks" and then removed

15   from the LIB. In the event that the number of significant coefficients is one, and the size of the block is one, this entry is a single coefficient and its coordinates are moved to LSP.

There are two methods available to process the sub-blocks. The first method, known

20   as depth-first quadtree coding, inserts the four sub-blocks into LIB immediately following the position of their parent block. The four child blocks are then evaluated immediately with respect to their significance and this operation is applied recursively until no more subdivision is possible. When all significant coefficients in this block are found and moved into LSP, the coding of the present entry is completed. The process

25   then moves to the next block in the LIB.

The second method, or breadth-first quadtree coding, adds these four sub-blocks to the end of LIB where they are evaluated before the same pass ends. With the breadth-first process, all parent squares at the same level will be processed before any blocks of the

30   next generation.

After all entries in the present LIB have been processed at one level of significance, the entries in TLIB are reordered according to the size of the block: each block must be put before those blocks with larger size so that it can be processed first for the next threshold. Most pixels adjacent to significant pixels have been moved into TLIB as

5   pixel level entries if not significant to the present threshold. Due to the correlation of adjacent coefficients, it is very likely that these adjacent pixels will be significant at the next threshold level. In the event of a strict bit budget, we must put these pixel level blocks first to ensure that precious bits are not used to find significant coefficients from big blocks, and risk missing pixel level significant coefficients. The reordering of TLIB

10   will therefore aid the encoding of more significant coefficients using fewer bits. While not essential, experiments show that higher PSNR will be achieved using this reordering scheme. The final step in this quadtree process is to replace the LIB with TLIB for subsequent scanning at the next level of significance and to reset TLIB to empty. Before moving to the next threshold however, the refinement data for

15   significant coefficients is collected.

Figure 7 illustrates the refinement pass, in the quadtree encoding of the image data. For those coefficient entries of LSP that are significant at threshold 2^{N+1} ($|c_{i,j}| \geq 2^{N+1}$), output its N-th bit. As illustrated in Figure 3 and discussed above, following the

20   refinement pass, the threshold is divided by 2 and the above process resumes with the new LIB --formerly the TLIB -- and the new threshold value.

The arithmetic coding of the bit stream produced by the above process is not essential. There are two types of data in the bit-stream: quadtree-coded significance map

25   encoding bits and refinement bits, which form a completely embedded code. There are many ways to organize this bitstream. In theory, the significance map data and the refinement bits data can be merged together in any order. This is handled by a multiplexer which packs the data according to user-specified priority. The default order of data packing is illustrated in Figure 9 and ensures optimum results when high

30   PSNR is pursued.

At the first stage of decoding, the following information must be reconstructed from the header bits: the starting threshold index N, the number of wavelet scales, and the image size. Based on the above information, we can initialize and fill LIB while the initial LSP and TLIB are set empty. The initial value of all wavelet coefficients is set

5   to zero.

The key process of decoding is illustrated in Figure 11, in which the significance map at a given threshold level is decoded based on the received bits. Assuming the present threshold index is N, the process first loads an entry from the LIB and reads one bit

10. from the bitstream. If the bit value is zero, this entry is moved to TLIB. In the alternative, the entry is checked to determine if its size is one. If the entry is a single pixel, then update the wavelet coefficient at the current position as $2^N + 2^{N-1}$, and read in one more bit. If this bit is a 1, update the coefficient at this position as $-(2^N + 2^{N-1})$. The entry is then moved into the LSP. If the entry is not at pixel level,

15   the process decomposes it into four equal sized sub-blocks. If the encoder has used the depth-first method (this decision having been made by the encoder and which information is contained in the header part of the bitstream), insert the sub-blocks into LIB at its parent block position. If the encoder has used the breadth-first method, add the sub-blocks to the end of LIB. After all entries in LIB have been decoded, using

20   TLIB to replace LIB, which will be processed at the next threshold level. The LIB is reordered according to the same rule as in encoding, and the TLIB is reset as empty.

In the refinement pass of the decoding, all coefficients, which have been moved into LSP, are updated according to following rule: if the coefficient is negative, then add

25   2^{N-1} if received bit is 0, or subtract 2^{N-1} if received bit is 1. On the contrary, if the coefficient is positive, then add 2^{N-1} if received bit is 1, or subtract 2^{N-1} if received bit is 0.

At any point in the encoding or decoding process of the present invention, bit

30   consumption may be calculated to determine if the bit budget has been exceeded and the process may be halted. In this manner, precise bit rate control can be easily

achieved if there is no arithmetic coding on the bit stream. With arithmetic coding, the

resultant bitstream is usually shorter than the desired length.